

# WINSTEAD



115er  
~~AF\$~~  
2100

October 6, 2003

Mail Stop Appeal Brief - Patents  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Docket No.: 7036-P145US

## RECEIVED

OCT 23 2003

Technology Center 2600

Appellant: John Brian Pickering  
Title: Voice Processing System  
Docket No.: GB9-1999-0107US1 (a/k/a 7036-P145US)  
Serial No.: 09/696,485 Filing Date: 10/25/2000  
Examiner: Susan Iris McFadden Group Art Unit: 2654

Transmitted herewith are the following:

1. Return Receipt Postcard
2. This Transmittal Letter (2 pages) (in duplicate)
3. 19 pages of Appeal Brief Under 37 C.F.R. § 1.191 (in triplicate)

for the above-identified Application.

- ☐ No additional fee is required.  
☒ The fee has been calculated as shown below:

CLAIMS AS AMENDED						
	CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE	ADDITIONAL FEE
Total Claims:	27	-	27	0	x \$18=	\$ 0.00
Independent Claims:	4	-	4	0	x \$84 =	\$ 0.00
TOTAL ADDITIONAL FEE FOR THIS AMENDMENT =						\$ 0.00

- ☒ Appeal Brief fee \$320.00  
☒ Please charge Deposit Account No. 50-0563 (7036-P145US)  
in the name of IBM Corporation in the amount of \$320.00

SUITE 800  
100 CONGRESS AVENUE  
AUSTIN, TEXAS 78701

PH 512.474.4330  
FAX 512.370.2850  
WINSTEAD.COM

WINSTEAD SECHREST & MINICK  
Attorneys and Counselors  
A Professional Corporation

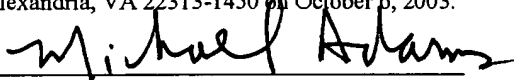
Austin, Dallas, Fort Worth,  
Houston, Mexico City,  
The Woodlands, Washington, DC



If the fee has been calculated incorrectly, the Commissioner is hereby authorized to charge any insufficiency of payment of the following fees associated with this communication, or credit any overpayment, to Deposit Account No. 50-0563 (7036-P145US) in the name of IBM Corporation. A duplicate copy of this transmittal letter is enclosed.

CERTIFICATION UNDER 37 C.F.R. § 1.8

I hereby certify that this correspondence (along with any item referred to as being enclosed herewith) is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to Mail Stop Appeal Brief - Patents, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on October 6, 2003.

  
Signature

Respectfully submitted,



Michael P. Adams  
Attorney for Applicant(s)  
Reg. No. 34,763

MPA:wk

Enclosure(s)



PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Appellant: John Brian Pickering  
Assignee: IBM Corporation  
Title: Voice Processing System

Serial No.: 09/696,485 Filing Date: 10/25/2000  
Examiner: Susan Iris McFadden Group Art Unit: 2654  
Docket No.: GB9-1999-0107US1 Conf. No.: 3603  
(a/k/a 7036-P145US)

RECEIVED

OCT 23 2003

Technology Center 2600

Dallas, Texas  
October 6, 2003

Mail Stop Appeal Brief - Patents  
COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, VA 22313-1450

**APPEAL BRIEF UNDER 37 C.F.R. § 1.191**

Dear Sir:

Appellant submits this Appeal Brief pursuant to the Notice of Appeal filed in this case on August 6, 2003.

The Commissioner is hereby authorized to deduct from Deposit Account No. 50-0563 in the name of IBM Corporation, the amount of \$320.00, being the amount specified in 37 C.F.R. 1.17(c) for this Appeal Brief. The Commissioner is also authorized to deduct any other amounts required for this Appeal Brief and to credit any amounts overpaid to Deposit Account No. 50-0563. This paper is submitted in triplicate.

**I. REAL PARTY IN INTEREST**

The real party in interest is the assignee, IBM Corporation, as named in the caption above.

10/21/2003 BLUANG1 00000008 500563 09696485  
01 FC:1402 330.00 DA

**II. RELATED APPEALS AND INTERFERENCES**

Based on information and belief, there are no appeals or interferences that could directly affect or be directly affected by or have a bearing on the decision by the Board of Patent Appeals in the pending appeal.

**III. STATUS OF CLAIMS**

Claims 1-27 are pending in the application. Claims 1-27 stand rejected.

**IV. STATUS OF AMENDMENTS**

The Appellant's response and claim amendments to the Office Action, having a mailing date of December 18, 2002, have been considered, but the Examiner indicated that they did not place the application in condition for allowance because the Appellant's arguments were deemed unpersuasive. The Appellant's response to the Final Office Action, having a mailing date of May 7, 2003, has also been considered, but the Examiner again indicated that it did not place the application in condition for allowance because the Appellant's arguments were deemed unpersuasive.

**V. SUMMARY OF THE INVENTION**

The present invention provides a method of providing speech recognition with barge-in for a voice processing system. The method plays out a prompt to a user, receives audio input from the user while the prompt is still being played out, performs speech recognition on the audio input to determine a corresponding text, performs lexical analysis on the text to determine whether or not the text satisfies one or more conditions, and responsive to the text satisfying the one or more conditions, terminates the play out of the prompt, otherwise it continues the play out of said prompt.

The voice processing system is allowed to assess whether or not to give effect to barge-in, and interrupt the outgoing prompt, based on an analysis of what has actually been said. In particular, this allows the system to discriminate some background comment, whether by the user or someone with

them, from a positive response to the prompt. Thus, where the input is determined to be not relevant to the prompt, the play out of the prompt is continued, and the user is not left accidentally suspended in an application script. In such a case, the recognized text can generally be discarded as irrelevant to the actual dialogue between the user and the voice processing system.

In one embodiment, the lexical analyzer determines whether or not the response is relevant to the prompt by scanning the text to see if it contains one or more predetermined words (note that the list of possible words to be matched may vary with each prompt). For instance, if a U.S. caller is asked to name the state in which they live, the lexical analyzer may simply be provided with a list of 50 states to match dynamically. Specification, page 6, line 3 through page 7 line 2.

As acknowledged by Appellant in the "Background Information" section of the application, barge-in systems have been known in the prior art. However, such prior art barge-in systems have limitations. For example, some barge-in systems may suspend the outgoing prompt upon receiving any response or input, including extraneous noise, even though the caller actually still desires to hear the rest of the prompt. Specification, page 4, lines 5-11. Another prior art approach suspends the outgoing prompt only if certain pre-defined and limited user inputs are recognized. However, such a system, known as a "small vocabulary" system, is only effective where the user's likely responses to the prompt are limited. Such a prior art system is not feasible for modern voice processing applications which involve large vocabulary speech recognition. Specification, page 4, line 12 through page 5, line 3.

## **VI. ISSUES**

A. Are claims 1-11, 13-24, and 26-27 properly rejected under 35 U.S.C. § 102(e) as being anticipated by U.S. Patent No. 6,311,159 to Van Tichelen, et al. (hereinafter "Van Tichelen")?

B. Are claims 12 and 25 properly rejected under 35 U.S.C. § 103(a) as being unpatentable over Van Tichelen in view of U.S. Patent No. 6,427,134 to Garner, et al. (hereinafter "Garner")?

**VII. GROUPING OF THE CLAIMS**

Claims 1, 7, 14, and 20 form a first group.

Claims 2, 8, 15, and 21 form a second group.

Claims 3, 9, 16, and 22 form a third group.

Claims 4, 10, 17, and 23 form a fourth group.

Claims 5, 11, 18, and 24 form a fifth group.

Claims 6, 13, 19, and 26 form a sixth group.

Claims 12 and 25 form a seventh group.

Claim 27 should not be grouped together with the other claims and should be considered separately.

The reasons for these groupings are set forth in Appellant's arguments in Section VIII.

**VIII. ARGUMENT**

A. Claims 1-11, 13-24, and 26-27 are not properly rejected under 35 U.S.C. § 102(e) as being anticipated by Van Tichelen.

The Examiner has rejected Claims 1-11, 13-24, and 26-27 under 35 U.S.C. § 102(e) as being anticipated by Van Tichelen. *See, e.g.,* Paper No. 5, page 3.

For a claim to be anticipated under 35 U.S.C. § 102, each and every claim limitation must be found within the cited prior art reference. M.P.E.P. § 2131. Appellant respectfully submits that not each and every claim limitation of Appellant's claimed invention is found within the Van Tichelen reference. Accordingly, Appellant respectfully traverses these claim rejections.

Van Tichelen teaches a speech controlled computer user interface for managing communications between a user and one or more computer applications.

The user interface has a speech layer, an utterance layer, and a discourse layer. The speech layer is in communication with the user and converts between speech messages and text messages. The utterance layer is in communication with the speech layer, and converts between text messages and semantic meaning messages. The discourse layer is in communication with the utterance layer and the at least one application program, and processes messages from the user and the at least one application program, and generates responsive messages to the user and the at least one application program.

See Abstract of Van Tichelen.

Van Tichelen also discloses that it supports barge-in capability. *See, e.g.,* Col. 14, lines 55-62. However, the Examiner has not cited to any teachings which reflect that Van Tichelen teaches barge-in with the capabilities set forth in Appellant's claimed invention. The disclosure concerning barge-in capability in Van Tichelen is very limited, and there do not appear to be any teachings which reflect that the barge-in capability has any more functionality than what was known in the prior art described above. For example, barge-in capability of Van Tichelen may be limited to a system which interrupts the prompt upon receiving any user input, including non-responsive noise. Van Tichelen may also only support barge-in for "small vocabulary" applications. The Examiner has not offered any evidence which reflects that Van Tichelen teaches barge-in with any more capabilities than these prior art implementations.

As stated above, the teachings in Van Tichelen relating to barge-in functionality are very limited. However, based on what is disclosed, it appears that Van Tichelen actually teaches away from barge-in capability which terminates a prompt based on "*text satisfying one or more conditions*," as set forth in Appellant's claimed invention. For example, Van Tichelen refers to conversation management as being both "event driven" and "data driven." Van Tichelen then states: "Being event driven supports features such as . . . barge-in capability, time outs . . . . From a data driven perspective, conversation data 703 uses data frames and slots with verification and confirmation." Col. 14, lines 55-60 (emphasis added). "Conversation data 703 related to open instances of dialogue is maintained. Each user conversation is modeled as a sequence of several smaller dialogues. As depicted in Fig. 7, this conversation data 703 is kept in the form of data frames 704 for each dialogue which are managed in a frame stack from which one dialogue may call another. Each dialogue data frame 704 has various specified slots in which relevant data is maintained with slot values being in the form of lists with an associated belief strength – unknown, ambiguous, unconfirmed, confirmed, etc." Col. 14, lines 28-37 (emphasis added). Thus, from these teachings, Van Tichelen appears to indicate that its barge-in capability only occurs based on some event occurring (*e.g.*, any input or response from a user) and is not data driven (based on an analysis of the actual content of the user's conversation).

With regard to prompts, Van Tichelen merely teaches that they may be modal (a dialogue cannot continue without input) or non-modal (a dialogue can continue without input). See Col. 15, lines 4-8. Van Tichelen does not disclose that the input has to satisfy any particular conditions in order for the dialogue to continue in response to a modal prompt. Van Tichelen also does not disclose having the playing out of the prompt suspended if certain conditions are satisfied. Having a non-specified input is the only requirement disclosed for the dialogue to continue; suspension of the



prompt before it is finished playing based on particular input conditions being satisfied is not discussed.

Thus, Van Tichelen does not teach each and every limitation of Appellant's claimed invention. For example, Van Tichelen does not teach "*performing speech recognition on said audio input to determine a corresponding text*" and then "*terminating the playing out of the prompt*" based on "*said text satisfying said one or more conditions, otherwise, continuing the playing out of said prompt.*" In fact, Van Tichelen is directed to having a user begin a dialogue after the system is initialized; it is not directed to having the user respond to a prompt. *See, e.g.*, Col. 12, line 64 – Col. 13, line 40, where the user calls into an e-mail application by telephone and asks the system for new e-mail. Furthermore, as discussed above, Van Tichelen does not teach suspension of the prompt before it is finished playing based on particular input conditions being satisfied.

In the Final Office Action, the Examiner asserts without support or a cite to teachings in prior art that "one of ordinary skill in the art familiar with barge-in systems know that systems that accurately recognize words which satisfy desired conditions turn off the prompt being generated." Paper No. 5, page 2. The Examiner did not offer any objective evidence, such as cites to prior art, in support of this conclusory assertion. Accordingly, under the Federal Circuit case law discussed below, the Examiner's argument does not adequately support a determination of unpatentability.

In a recent decision from the United States Court of Appeals for the Federal Circuit, the Federal Circuit noted that when the patent examiner and Board "rely on what they assert to be general knowledge to negate patentability, that knowledge must be articulated and placed on the record." *In re Sang-Su Lee*, 277 F.3d 1338, 1345 (Fed. Cir. 2002). Specifically, the Federal Circuit noted that conclusory statements about what is "basic knowledge" or "common sense" by themselves do not adequately support a determination of unpatentability. *See id.* at 1343-44.

In the Examiner's Office Actions and Advisory Action, the Examiner also indicates that particular elements of Appellant's claimed invention are inherently taught in the Van Tichelen reference. For example, in the Advisory Action, the Examiner states that: "Automatic Speech Recognition systems inherently perform speech recognition on a user's response and compare this to allowable answers." Paper No. 7, page 2. *See also* Paper No. 5, page 3: "In regard to claims 2, 8, 15, and 21, the step of discarding said text is inherent when the internal timers time out (Col. 15, lines 5-11)." Appellant respectfully submits that the Examiner has not provided sufficient rationale or evidence necessary to show inherency, as required by M.P.E.P. § 2112. As stated in M.P.E.P. § 2112:

To establish inherency, the intrinsic evidence 'must make clear that the missing descriptive matter is necessarily present in the thing described in the reference, and that it would be so recognized by persons of ordinary skill. Inherency, however, may not be established by probabilities or possibilities. The mere fact that a certain thing may result from a given set of circumstances is not sufficient.' *In re Robertson*, 169 F.3d 743, 745, 49 USPQ2d 1949, 1950-51 (Fed. Cir. 1999) (citations omitted) . . . . In relying upon the theory of inherency, the examiner must provide a basis in fact and/or technical reasoning to reasonably support the determination that the allegedly inherent characteristic necessarily flows from the teachings of the applied prior art. *Ex parte Levy*, 17 USPQ2d 1461, 1464 (Bd. Pat. App. & Inter. 1990) (emphasis in original).

*See* M.P.E.P. § 2112 ("Examiner Must Provide Rationale or Evidence Intending to Show Inherency.")

Appellant respectfully submits that the Examiner has not offered any extrinsic evidence, basis in fact, and/or technical reasoning to reasonably support the determination that any alleged inherent characteristics of Appellant's claimed invention necessarily flow from the teachings of Van Tichelen reference.

Finally, although the Examiner has also rejected claim 27 in view of the Van Tichelen reference. The Examiner has not provided any support for this rejection. Claim 27, which is dependent on Claim 1, includes the additional limitation of: *"determining if said audio input is speech input, and wherein if said audio input is speech input, said step of performing speech recognition comprises performing speech recognition on said speech input to determine a corresponding text."* The Examiner did provide support for the argument that the limitation in claim 1 of *"performing speech recognition on said audio input to determine a corresponding text"* can be found in the Van Tichelen reference. However, the Examiner did not provide any support, such as cites to Van Tichelen, for the contention that the additional limitation of claim 27 is found in that reference. *See, e.g.*, Paper No. 5, page 3 (in rejecting claims 1, 7, 14, 20, and 27, the Examiner cites to Col. 1, Fig. 4 of Van Tichelen as teaching "circuitry that performs speech recognition on the audio input to determine a corresponding text," but does not provide any cites to Van Tichelen or any other support for the argument that this reference teaches *"determining if said audio input is speech input"*). Accordingly, this is an additional basis that the Examiner's rejection of Claim 27 is unfounded.

**B. Claims 12 and 25 are not properly rejected under 35 U.S.C. § 103(a) as being unpatentable over Van Tichelen in view of Garner.**

The Examiner further rejected claims 12 and 25 under 35 U.S.C. 103(a) as being unpatentable over Van Tichelen in view of U.S. Patent No. 6,427,134 to Garner ("Garner"). Appellant submits that claims 12 and 25 are allowable over these patents for at least the same reasons discussed above. Furthermore, Appellant also respectfully submits that the Examiner has not met the Examiner's burden of factually supporting the alleged motivation to combine the two patents.

It is the Examiner's burden to factually support any prima facie conclusion of obviousness. The Examiner's duty may not be satisfied by engaging in impermissible hindsight; any conclusion of obviousness must be reached on the basis of facts gleaned from the prior art. The preferred evidence

to be offered by the Examiner is an express teaching to modify/combine which is set forth within objectively verifiable sources of prior art. See M.P.E.P. §§ 2141-2144. Appellant respectfully submits that the Examiner has not satisfied the burden of factually supporting the alleged motivation to combine the Van Tichelen and Garner patents. For example, the Examiner has not cited to any express teachings within these patents which support a motivation to combine these patents to achieve Appellant's claimed invention. Therefore, Appellant submits that the Examiner has not established a prima facie case of obviousness, and Claims 12 and 25 should therefore be allowed.

More particularly, Appellant submits that the Examiner has not cited to any teachings that would support a motivation to combine the voice activity detector especially suitable for use in mobile phones which may be required to operate in noisy environments, as taught in Garner, with the speech controlled computer user interface which communicates between a user and an application program, as taught by Van Tichelen. Garner's voice activity detector distinguishes between voice and speech, and if the signal is noise it is not transmitted. *See Abstract.* Garner does not distinguish between noise and speech to determine "*whether to continue or terminate playing out of said prompt.*" Thus, Appellant submits that the Examiner has: (1) not cited to any teachings that would support a motivation to combine the teachings of Garner and Van Tichelen; and (2) not cited to teachings that would achieve all of the elements of Appellant's claimed invention.

#### **IX. CONCLUSION**

For the above reasons, Appellant respectfully submits that rejection of pending Claims 1-27 is unfounded. Accordingly, Appellant requests that the rejection of Claims 1-27 be reversed.

This Brief is submitted in triplicate.

Respectfully submitted,

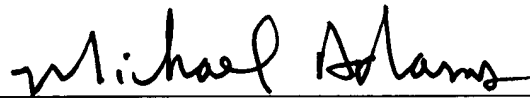


PATENT

Michael P. Adams  
Attorney for Appellant(s)  
Reg. No. 34,763

**CERTIFICATION UNDER 37 C.F.R. § 1.8**

I hereby certify that this correspondence (along with any item referred to as being enclosed herewith) is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to Mail Stop Appeal Brief - Patents, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on October 6, 2003.



Signature

**APPENDIX**

1. A method of providing speech recognition with barge-in for a voice processing system comprising the steps of:

playing out a prompt to a user;

receiving audio input from the user while said prompt is still being played out;

performing speech recognition on said audio input to determine a corresponding text;

performing lexical analysis on said text to determine whether the text satisfies one or more conditions which reflect a relevant user response to said prompt; and

responsive to said text satisfying said one or more conditions, terminating the playing out of the prompt;

otherwise, continuing the playing out of said prompt.

2. The method of claim 1, further comprising the step, responsive to said step of continuing the playing out of said prompt, of discarding said text.

3. The method of claim 1, wherein said step of performing lexical analysis to determine whether the text satisfies one or more conditions comprises the step of scanning the text to see if it contains one or more predetermined words.

4. The method of claim 3, wherein said one or more predetermined words are specific to the particular prompt being played out.

5. The method of claim 1, in which said voice processing system and said user communicate with each other over a telephone network, whereby the prompt is played out over a telephone connection, and said audio input is received back over the telephone connection.

6. The method of claim 1, further comprising the step of using one or more acoustic parameters of the audio input to assist determining whether to continue or to terminate playing out of said prompt.

7. A voice processing system for providing speech recognition with barge-in, said voice processing system comprising:

means for playing out a prompt to a user;

means for receiving audio input from the user while said prompt is still being played out;

means for performing speech recognition on said audio input to determine a corresponding text;

means for performing lexical analysis on said text to determine whether the text satisfies one or more conditions which reflect a relevant user response to said prompt;

means responsive to said text satisfying said one or more conditions, for terminating the playing out of the prompt and;

otherwise, means for continuing the playing out of said prompt.

8. The voice processing system of claim 7, further comprising the means, responsive to said means of continuing the playing out of said prompt, for discarding said text.

9. The voice processing system of claim 7, wherein said means for performing lexical analysis to determine whether the text satisfies one or more conditions comprises means for scanning the text to see if it contains one or more predetermined words.

10. The voice processing system of claim 9, wherein said one or more predetermined words are specific to the particular prompt being played out.



11. The voice processing system of claim 7, in which said voice processing system and said user communicate with each other over a telephone network, whereby the prompt is played out over a telephone connection, and said audio input is received back over the telephone connection.

12. The voice processing system of claim 7, wherein said means for receiving caller input includes a voice activity detector for discriminating between speech input and other forms of tone or noise input.

13. The voice processing system of claim 7, further comprising means for calculating one or more acoustic parameters of the audio input to assist determining whether to continue or to terminate playing out of said prompt.

14. A computer readable medium containing computer program instructions for a voice processing system for providing speech recognition with barge-in, said computer program instructions comprising instructions for:

playing out a prompt to a user;

receiving audio input from the user while said prompt is still being played out;

performing speech recognition on said audio input to determine a corresponding text;

performing lexical analysis on said text to determine whether the text satisfies one or more conditions which reflect a relevant user response to said prompt;

responsive to said text satisfying said one or more conditions, terminating the playing out of the prompt and;

otherwise, continuing the playing out of said prompt.

15. The computer readable medium of claim 4, further comprising the instruction, responsive to said instruction of continuing the playing out of said prompt, of discarding said text.

16. The computer readable medium of claim 4, wherein said instruction of performing lexical analysis to determine whether the text satisfies one or more conditions comprises the instruction of scanning the text to see if it contains one or more predetermined words.

17. The computer readable medium of claim 16, wherein said one or more predetermined words are specific to the particular prompt being played out.

18. The computer readable medium of claim 4, in which said voice processing system and said user communicate with each other over a telephone network, whereby the prompt is played out over a telephone connection, and said audio input is received back over the telephone connection.

19. The computer readable medium of claim 4, further comprising the instruction of using one or more acoustic parameters of the audio input to assist determining whether to continue or to terminate playing out of said prompt.

20. A voice processing system for providing speech recognition with barge-in, said voice processing system comprising:

circuitry for playing out a prompt to a user;

circuitry for receiving audio input from the user while said prompt is still being played out;

circuitry for performing speech recognition on said audio input to determine a corresponding text;

circuitry for performing lexical analysis on said text to determine whether the text satisfies one or more conditions which reflect a relevant user response to said prompt;

circuitry responsive to said text satisfying said one or more conditions, for terminating the playing out of the prompt and;

otherwise, circuitry for continuing the playing out of said prompt.

21. The voice processing system of claim 20, further comprising circuitry, responsive to said circuitry of continuing the playing out of said prompt, for discarding said text.

22. The voice processing system of claim 21, wherein said circuitry for performing lexical analysis to determine whether the text satisfies one or more conditions comprises circuitry for scanning the text to see if it contains one or more predetermined words.

23. The voice processing system of claim 22, wherein said one or more predetermined words are specific to the particular prompt being played out.

24. The voice processing system of claim 23, in which said voice processing system and said user communicate with each other over a telephone network, whereby the prompt is played out over a telephone connection, and said audio input is received back over the telephone connection.

25. The voice processing system of claim 24, wherein said circuitry for receiving caller input includes a voice activity detector for discriminating between speech input and other forms of tone or noise input.

26. The voice processing system of claim 25, further comprising circuitry for calculating one or more acoustic parameters of the audio input to assist determining whether to continue or to terminate playing out of said prompt.

27. The method of claim 1, further comprising the step of determining if said audio input is speech input, and wherein if said audio input is speech input, said step of performing speech recognition comprises performing speech recognition on said speech input to determine a corresponding text.

AUSTIN 1\232926\1  
7036-P145US 10/06/2003